

STATINTL

November 3, 1952

Mr. James Q. Reber
2430 E Street, N.W.
Washington 25, D. C.

Dear Jim:

This will confirm my advices to you on the phone to the effect (1) that [] will visit the Agency on Wednesday, November 5, his agenda to be in [] hands and (2) that [] will appear on Thursday, November 6, for discussions on machine indexing and machine translation.

My suggested agenda for them is (1) a meeting with you to clarify the present position up here with respect to these projects and their idea about possible extension; (2) a meeting with those primarily interested in machine indexing which should include []
OCD, somebody from OCI, somebody from ORR (possibly []
and [] from Management; (3) a meeting on machine translation with [] and anyone else that you or he deems appropriate.

My view about this set of activities is as follows. I believe that the problem of indexing information so that it can be searched for and recovered with a maximum of precision in a minimum of time is one of the most if not the most important problem facing the Intelligence community. Time and again in my work done there I ran into situations in which we discovered by accident that there was material directly bearing on problems we were working on which we had no systematic way of knowing was there. The importance of the machine aspect of this problem is not that the machines can in any way substitute for analytical work. Of course, you get out of a machine no more than you put into it. On the other hand, the effectiveness of any indexing system is measured by the speed and precision with which the analyst can secure a comprehensive set of references to all the available evidence that might bear on his problem. I am convinced that machine methods can vastly increase the efficiency

Mr. James G. Reber

2

November 3, 1952

STATINTL

STATINTL

of reference services, and I think Perry and his group have made real advances in the design of an appropriate system. I would, therefore, put very great pressure indeed on the Agency not to drop this whole effort at this stage. This, of course, is a different question from the question of whether or not it is appropriate for some part of the work to be carried on in [] for a longer period than originally contemplated. On this I have very little judgment, though I am much impressed by the arguments that [] advances for at least another six month's work beyond the present January 1 deadline.

I also feel that the machine indexing problem should not be looked at separately from the machine translation problem. There is a major problem in the whole field of foreign documents of indexing foreign language materials in advance of translation so that as problems arise researchers can determine whether there is foreign language material bearing on those problems which deserves to be translated. I think you and I are in essential agreement on the appropriate role of FDD in this activity. FDD has not been much interested in the machine translation problem because, given their present philosophy, - that their staff consists of a group of analysts who will select, abstract, and translate all in one operation, - there is very little need for a translating machine. If as I believe should be done, their function is redefined as primarily that of insuring that foreign language materials are so processed that analysts can recover relevant information from them, the relevance being determined by the analyst and not the translator then I think there is a really big prospect in a combination of machine indexing and machine translation. It should be possible to develop a system under which titles, subject matter heading, and other elements of an indexing code could be punched into a machine directly in the original language but in which searching by the analyst could be carried on under English language subject headings.

I raise this possibility because I think neither [] nor those previously interested in machine indexing have yet seen the possibility of this marriage of the two techniques. For fruitful work to be done on this kind of combined system, one needs people of three sorts of talents (1) people who know something about indexing, (2) people who know something about language and (3) people who know something about machines.

STATINTL

Mr. James G. Reber

3

November 3, 1952

STATINTL

I think [] is probably right in feeling that we need some help from others on the language side, but I think he underestimates the purely machine problems on which [] people are pre-eminent and so far as I know he has never given any thought to the indexing problems.

STATINT

I believe we have enough leeway in the [] budget to provide an additional twenty to twenty-five thousand dollars toward this enterprise from our existing contract if the Agency should regard this as an appropriate use of the funds. Anything more than this would have to be separately appropriated in addition to our present contract.

STATINT

Best regards.

Yours,

[] (Signed)

STATINT